BISS Biodiversity Information Science and Standards

Biodiversity Information Science and Standards 5: e74275
doi: 10.3897/biss.5.74275

OPEN ACCESS

Conference Abstract

Unlocking Inventory Data Capture, Sharing and Reuse: The Humboldt Extension to Darwin Core

Yanina V. Sica[‡], Paula F. Zermoglio[§,|]

‡ Department of Ecology and Evolutionary Biology, Center for Biodiversity and Global Change, Yale University, New Haven, United States of America
§ Instituto de Ecología, Genética y Evolución (IEGEBA-CONICET), Buenos Aires, Argentina
| VertNet, Bariloche, Argentina

Corresponding author: Yanina V. Sica (yanina.sica@gmail.com)

Received: 10 Sep 2021 | Published: 13 Sep 2021

Citation: Sica YV, Zermoglio PF (2021) Unlocking Inventory Data Capture, Sharing and Reuse: The Humboldt Extension to Darwin Core. Biodiversity Information Science and Standards 5: e74275.
https://doi.org/10.3897/biss.5.74275

Abstract

Biodiversity inventories, i.e., recording multiple species at a specific place and time, are routinely performed and offer high-quality data for characterizing biodiversity and its change. Digitization, sharing and reuse of incidental point records (i.e., records that are not readily associated with systematic sampling or monitoring, typically museum specimens and many observations from citizen science projects) has been the focus for many years in the biodiversity data community. Only more recently, attention has been directed towards mobilizing data from both new and longstanding inventories and monitoring efforts. These kinds of studies provide very rich data that can enable inferences about species absence, but their reliability depends on the methodology implemented, the survey effort and completeness. The information about these elements has often been regarded as metadata and captured in an unstructured manner, thus making their full use very challenging.

Unlocking and integrating inventory data requires data standards that can facilitate capture and sharing of data with the appropriate depth. The Darwin Core standard (Wieczorek et al. 2012) currently enables reporting some of the information contained in inventories, particularly using Darwin Core Event terms such as samplingProtocol, sampleSizeValue, sampleSizeUnit, samplingEffort. However, it is limited in its ability to accommodate spatial,

© Sica Y, Zermoglio P. This is an open access article distributed under the terms of the Creative Commons Attribution License (CC BY 4.0), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

temporal, and taxonomic scopes, and other key aspects of the inventory sampling process, such as direct or inferred measures of sampling effort and completeness. The lack of a standardized way to share inventory data has hindered their mobilization, integration, and broad reuse. In an effort to overcome these limitations, a framework was developed to standardize inventory data reporting: Humboldt Core (Guralnick et al. 2018). Humboldt Core identified three types of inventories (single, elementary, and summary inventories) and proposed a series of terms to report their content. These terms were organized in six categories: dataset and identification; geospatial and habitat scope; temporal scope; taxonomic scope; methodology description; and completeness and effort. While originally planned as a new TDWG standard and being currently implemented in Map of Life (https://mol.org/humboldtcore/), ratification was not pursued at the time, thus limiting broader community adoption.

In 2021 the TDWG Humboldt Core Task Group was established to review how to best integrate the terms proposed in the original publication with existing standards and implementation schemas. The first goal of the task group was to determine whether a new, separate standard was needed or if an extension to Darwin Core could accommodate the terms necessary to describe the relevant information elements. Since the different types of inventories can be thought of as Events with different nesting levels (events within events, e.g., plots within sites), and after an initial mapping to existing Darwin Core terms, it was deemed appropriate to start from a Darwin Core Event Core and build an extension to include Humboldt Core terms. The task group members are currently revising all original Humboldt Core terms, reformulating definitions, comments, and examples, and discarding or adding new terms where needed. We are also gathering real datasets to test the use of the extension once an initial list of revised terms is ready, before undergoing a public review period as established by the TDWG process.

Through the ratification of Humboldt Core as a TDWG extension, we expect to provide the community with a solution to share and use inventory data, which improves biodiversity data discoverability, interoperability and reuse while lowering the reporting burden at different levels (data collection, integration and sharing).

Keywords

biodiversity surveys, standardization, event

Presenting author

Yanina V. Sica

Presented at

TDWG 2021

Acknowledgements

TDWG Humboldt Core Task Group members

References

- Guralnick R, Walls R, Jetz W (2018) Humboldt Core toward a standardized capture of biological inventories for biodiversity monitoring, modeling and assessment. Ecography 41: 713-725. https://doi.org/10.1111/ecog.02942
- Wieczorek J, Bloom D, Guralnick R, Blum S, Dring M, Giovanni R, (2012) Darwin Core: An Evolving Community-Developed Biodiversity Data Standard. PLOS One 7 (1). https://doi.org/10.1371/journal.pone.0029715